

다중 시기/센서 원격탐사 자료의 통계비 기반 분류

Statistical ratio based classification of multi-temporal/sensor remote sensing data

박노옥*, 지광훈

No-Wook Park* and Kwang-Hoon Chi

한국지질자원연구원 지질자원정보센터

E-mail: nwpark@kigam.re.kr

요약

이 연구에서는 다중 센서 융합과 시간적 문맥 정보의 결합을 통한 분류 정확도 향상을 목적으로 통계비 기반 결정수준 융합 기법을 제안하였다. 다중 센서 융합을 목적으로 개별 센서 자료로부터 얻어진 사후 확률의 결합에 기존 확률론적 자료 융합에서 널리 사용되어온 조건부 독립의 가정을 완화한 통계비 기반 결합 규칙을 적용하였다. 그리고 시간적 문맥 정보를 새로운 정보 근원으로 간주하고 이전 시기 자료의 분류결과로부터 추출 및 결합하였다. 이 제안기법은 통계비 기반의 틀 안에서 다중 센서의 분광정보 및 시간적 문맥 정보의 결합이 용이한 장점이 있다. 제안기법의 적용성 평가를 위해 다중 시기/센서 융합의 사례연구를 수행하였다.

1. 서론

1990년대부터 다양한 센서 자료의 이용이 증가됨에 따라 자료 융합에 대한 연구가 많이 진행되고 있다. 이러한 자료 융합을 위해 기존 단일 센서 분석기법을 직접적으로 적용하기에는 한계가 있기 때문에 새로운 분석 방법론을 필요로 한다.

여러 자료 융합 방법론 중에서 베이저안 확률이론은 초기확률을 다양한 자료로부터 얻어지는 조건부 확률 혹은 사후 확률과 함께 결합하게 된다. 일반적으로 다양한 자료들에 대한 결합 확률모델의 정

의가 쉽지 않기 때문에, 계산상의 편의를 위해 조건부 독립의 가정하에 개별자료들의 조건부 확률 혹은 사후 확률들의 곱으로 결합 확률을 계산하게 된다. 이러한 조건부 독립의 가정은 개념적인 가정을 일부 포함하기 때문에 통계학의 독립 테스트로 규명하기에는 쉽지 않다. 또한 일부 자료에 따라서는 일반적인 확률의 조건 즉 0에서 1 사이의 값을 가진다는 조건을 위반하는 경우도 나타나게 된다.

이러한 문제에 대한 대안으로 Journal(2002)는 기존 조건부 독립 혹은 완전 독립의 가정하에 결합을 수행하는

대신에 증분 자체보다는 일반적으로 더 안정적인 증분의 비율 즉 통계비의 결합을 제안하였다. 그러나 위 논문이 이론적인 근거를 제시하였으나, 실제 자료에 대한 적용은 현재까지 진행되지 못한 상태이다. 또한 이론 자체를 일반적인 지구과학 현상에서 다루는 single-class 경우에 대해서 전개하여 multi-class를 다루는 원격탐사 화상 분야에 대한 확장은 이루어지지 않은 상태이다.

이 논문에서는 Journal(2000)에 의해 제안된 통계비 기반 자료 융합 규칙을 다중 센서 원격탐사 자료의 융합에 확장 적용하였다. 다중 시기/센서 융합의 경우에 적합하도록 융합 규칙을 적용하였으며, 사례 연구를 통해 적용성을 살펴보았다.

2. 방법론

(1) 통계비 기반 융합

통계비 기반 융합은 기존 조건부 독립 혹은 완전 독립의 가정 하에 결합을 수행하는 베이저안 이론의 대안으로 제시된 방법론이다.

토지 피복 구분을 위해 특정 토지 피복 A의 처리를 위해 우선 2개의 자료 B와 C를 고려하였을 때, 이 방법은 증분 자체보다는 일반적으로 더 안정적인 증분의 비율을 기본적으로 이용한다. A의 사전 확률, 자료 B와 C로부터 얻어진 사후 확률을 각각 $P(A)$, $P(A|B)$ 및 $P(A|C)$ 라고 하였을 때, 로지스틱 타입의 확률비인 a , b , c 및 이들의 결합인 x 를 다음과 같이 계산하게 된다(Journal, 2002).

$$a = \frac{1 - P(A)}{P(A)} \in [0, \infty],$$

$$b = \frac{1 - P(A|B)}{P(A|B)}, \quad c = \frac{1 - P(A|C)}{P(A|C)},$$

$$x = \frac{1 - P(A|B, C)}{P(A|B, C)}$$

위 식에서 a 는 A에 대한 사전 불확실성의 측정 혹은 A 발생에 대한 거리 개념으로 간주할 수 있다. 즉 만약 A가 확실하게 일어난다면 a 는 0이 되며, 발생이 불가능하면 ∞ 의 값을 가지게 된다. 비슷하게 c 는 자료 C를 관측한 후에 A 발생에 대한 거리 개념으로 간주할 수 있다. 이들의 비율인 $\frac{c}{a}$ 는 사전 거리 a 로부터 출발하여 자료 C의 감소 혹은 증가와 같은 증분의 기여도로 간주할 수 있다. x 는 자료 B와 C를 모두 관측한 후에 A의 발생에 대한 거리를 나타낸다.

이 방법에서는 이러한 비율에 기반하여 자료 C의 A 발생에 대한 증분의 기여도는 자료 B를 관측하기 전이나 후에 똑같다는 가정 하에 다음 식을 이용하게 된다

$$\frac{x}{b} \simeq \frac{c}{a}$$

이로부터 최종적으로 구하고자 하는 A의 B와 C에 대한 결합 사후 확률은 다음 식으로 계산되어 질 수 있다.

$$P(A|B, C) = \frac{1}{1+x} = \frac{a}{a+bc} \in [0, 1]$$

위의 관계식은 자료 사이의 확률인 $P(B, C)$ 를 필요로 하지 않기 때문에, 자료간의 종속 혹은 독립에 대한 가정을 필요로 하지 않는다.

이러한 과정을 n 개의 자료 D_i ($i=1, \dots, n$)에 일반화하면 다음과 같다

$$P(A|D_i, i=1, \dots, n) = \frac{1}{1+x}$$

$$x = \frac{\prod_{i=1}^n d_i}{a^{n-1}}, \quad a = \frac{1 - P(A)}{P(A)},$$

$$d_i = \frac{1 - P(A|D_i)}{P(A|D_i)}, \quad i = 1, \dots, n$$

(2) 다중 시기/센서 원격탐사 자료 융합

원격탐사 자료의 분류를 목적으로 이러한 다중 센서 자료의 이용뿐만 아니라, 부가적인 정보를 추출하여 이를 결합함으로써 분류 정확도를 향상시키려는 노력도 함께 진행되고 있다. 분류 정확도 향상을 위해 주변 화소와의 공간적인 상관성을 고려하는 공간적 문맥 분류는 많이 연구가 진행되어 왔으나, 시간적인 상관성을 고려하는 시간적 문맥 분류는 상대적으로 많은 연구가 진행되어 있지 않다. 특히 광학자료의 이용이 불가능한 상황에서 이용이 가능한 SAR 자료의 경우, 광학자료에 비해 높은 분류 정확도를 얻기가 쉽지 않은 상황이다. 분류를 목적으로 SAR 자료를 이용할 경우, 이전 시기에 얻어진 광학자료와의 시간적 상관성을 고려할 경우 분류 정확도를 향상시킬 가능성이 크다. 그동안 다중 센서 자료의 융합과 시공간적 문맥 분류를 목적으로 Markov Random Fields 등이 적용되어 상대적으로 높은 분류 정확도를 나타내지만, 각 항목별 기여도를 결정해야 하는 번거로움이 있다.

이 연구에서는 다중 센서로부터 추출되는 분광 혹은 산란 정보와 시간적 문맥 정보를 결합하였다. 기존 연구에서는 시간적 문맥 정보를 고려하는 두 시기간 토지 피복 항목간의 정성적 해석에 따라 주관적으로 할당하였으나, 이 연구에서는 이전 시기의 분류결과와 이후 시기의 훈련집단으로부터 추출하였다. 이전 시기의 자료를 일단 분류한 후에, 이 결과를 이후 시기의 훈련집단과 비교하여 정규 빈도수를 계산하도록 하였다. 이로부터 이전 시기 분류 항목별로 이후 시기 분류 항목에 대한 확률값을 얻게 된다. 이러한 확률값을 시간적 문맥 정보로 간주하고 이후 시기의 분

광 혹은 산란정보와 결합하였다.

3. 사례연구

제안 방법의 적용성 평가를 위해 다중 주파수/편광 인공위성 SAR 자료인 ERS, Radarsat-1, JERS-1 SAR를 이용한 분류 과정에 이전 시기에 획득된 광학 자료를 이용한 사례연구를 수행하였다.

다중 시기로 1998년 4월과 7월초를 고려하였는데, 시간적 문맥정보와 결합한 7월초 자료의 분류 정확도 향상 여부에 목적이 있다. 4월 자료로는 4월 19일에 촬영된 JERS-1 OPS 자료를 이용하였으며, 7월초의 자료로는 7월 5일에 촬영된 C 밴드 VV 편광상태의 ERS-2 SAR, 7월 7일에 촬영된 L 밴드 HH 편광상태의 JERS-1 SAR, 7월 9일에 촬영된 C 밴드 HH 편광상태의 Radarsat-1 자료를 이용하였다.

지형도, 현지조사와 광학 자료를 통해 4월과 7월의 훈련집단과 참조집단을 각각 작성하였다. 4월에는 논, 밭, 산림, 물, 도심, 비닐하우스, 보리, 객토, 나지 등의 9개 토지 피복 항목을 고려하였으며, 7월에는 논, 물이 차있는 논, 밭, 산림, 물, 도심, 비닐하우스, 나지 등의 8개 항목을 고려하였다. 시기적인 차이로 인해 4월에 보리와 객토였던 지역은 7월에는 물이 차있는 논과 벼가 자란 논으로 바뀌게 된다. 이러한 시기적 차이를 시간적 문맥 정보로 활용할 경우 분류 정확도의 향상을 기대할 수 있다.

우선 4월의 JERS-1 OPS 자료의 분류에는 최대우도분류기법을 적용하였다. 7월의 SAR 자료의 분류를 위해 각 자료별로 후방산란계수 이외에 GLCM(Gray Level Cooccurrence Matrix)로부터 homogeneity, contrast, standard deviation 및

entropy를 추출하여 후방산란계수를 포함하여 총 5개의 특징을 분류에 이용하였다. SAR 자료의 분류에는 정규분포 가정을 따르지 않는 자료의 특성을 고려하여 비모수적 방법인 역전파 알고리즘 기반 인공신경망 기법을 적용하여 사후 확률을 얻었다.

Table 1. Classification accuracy

Case	Overall accuracy	Kappa
ERS-2	50.88	0.40
Radarsat-1	50.14	0.40
JERS-1	40.02	0.29
ERS-2 + JERS-1	62.09	0.54
ERS-2 + Radarsat-1	61.85	0.54
Radarsat-1 + JERS-1	57.40	0.49
all	68.80	0.62
all + temporal contextual informaiton	92.31	0.90

시간적 문맥정보를 고려하지 않은 7월 자료의 분류 정확도를 살펴보면(Table 1), 개별 SAR 자료만을 이용하였을 때에 비해서 다중 주파수 혹은 다편광 자료를 이용하였을 때 분류 정확도가 향상되는 것을 확인할 수 있다. 그러나 전반적인 정확도는 광학 자료의 분류에서 기대할 수 있는 정도의 정확도만큼 높게 나오지는 않았다.

4월의 JERS-1 OPS 자료의 분류결과와 7월의 훈련집단의 비교를 통해 빈도비로부터 시간적 변이도 확률을 계산한 후에 이 값을 통계비 기반 융합 규칙에 새로운 정보로 결합하였다. 이 경우 전체정확도는 약 92%, 카파계수는 0.90으로 SAR 자료만을 이용하였을 때에 비해 두드러진 향상을 나타내었다. 이는 이전시기로부터 계산되는 시간적 문맥정보가 SAR 자료로부터

얻어지는 산란정보와 함께 결합됨으로써 SAR 자료만을 이용하였을 때 흔재되게 나타나는 토지피복항목간의 구분력이 향상되었기 때문인 것으로 판단된다.

4. 결론

다중 센서로부터 얻을 수 있는 분광 혹은 산란정보와 다중 시기 자료로부터 얻을 수 있는 시간적 문맥 정보의 효율적 결합을 위해 통계비 기반의 결정수준 자료 융합 방법론을 제안하고 사례연구를 수행하였다. 제안 기법의 적용 결과, 시간적 문맥 정보를 이용함으로써 기존 다중 센서 혹은 단일 센서만의 분광 혹은 산란 정보를 이용한 분류 결과보다 향상된 분류 정확도를 나타내었다. 특히 제안기법은 SAR 자료의 분류 정확도 향상을 위한 방법론으로도 적용이 가능하다.

이 연구에서 고려한 시간적 문맥정보는 이전 시기자료의 분류 결과에 기반하고 있기 때문에 기존 분류결과에서 발생하는 분류 오차 등이 전파될 가능성이 남아있다. 이를 해결하기 위해 상호 교차시기의 시간적 문맥정보를 연속적으로 적용할 필요가 있다. 또한 분광정보, 시간적 문맥정보 이외에 공간적 문맥정보를 함께 결합을 수행할 예정이다.

5. 참고문헌

Journal, A.G., 2002, Combining knowledge from diverse sources: an alternative to traditional data independence hypotheses, *Mathematical Geology*, 34(5): 573-596.